Chapter 5

Auditory Communication in Domestic Dogs: Vocal Signalling in the Extended Social Environment of a Companion Animal

Anna Magdalena Taylor, Victoria Frances Ratcliffe, Karen McComb, and David Reby
School of Psychology, University of Sussex, Falmer, UK

5.1 INTRODUCTION

Vocal signals play a key role in the communication systems of most mammalian species. Crucially, vocalisations can transmit information about the signaller that receivers may use to mediate their responses during sexual and social interactions. Amongst mammals, the vocal repertoire and perceptual abilities of the domestic dog *(Canis familiaris)* are particularly interesting because their communication system seems to have undergone significant changes as they have adapted to the human environment. Domestication is characterised by selective breeding, and whilst natural selection pressures are relaxed, human-controlled selection typically leads to the development of new traits (Price, 1999). An important aspect of human-driven selection in many domesticated species is a requirement for individuals to cooperate with people, which can lead to a differentiation of morphological and behavioural traits as compared to the wild ancestor. Because domestication tends to favour individuals that are able to exploit human perceptual abilities and biases (e.g., domestic cats: McComb et al., 2009) as well as those best able to perceive and make functional assessments of human vocal signals, this chapter considers not only dog–dog vocal communication, but also the use of vocal signalling in dog–human interactions.

In the first section of this chapter, we review existing knowledge on the vocalisations of domestic dogs with a view to understanding their communicative function. The different calls made by dogs are familiar to most humans; barks,

The Social Dog. http://dx.doi.org/10.1016/B978-0-12-407818-5.00005-X

growls, and whimpers are ubiquitous in our everyday environment and even non-dog owners are likely to form an opinion of what a dog might be signalling. Whilst the contextual use of these calls may provide us with a broad understanding of their functional role, in order to fully investigate the communicative potential of signals also, we need to look at their acoustic structure, which we characterise in this chapter using the source–filter framework of vocal production. We emphasise how the relationship between anatomical characteristics and acoustic output directly influences the type and reliability of information contained within a call. In this light, we discuss in detail how dog vocalisations are able to broadcast specific information about static characteristics of the signaller, such as body size, and dynamic attributes, such as motivational state. We also review evidence that receivers, including dogs, humans, and potentially other species, can perceive these acoustic cues to make functional assessments and adapt their own behavioural responses.

Throughout this chapter, it is evident that whilst domestic dog vocalisations retain socially relevant information that can be identified by other dogs, functional changes in their vocal communication system may also have occurred to facilitate their interaction with humans. Finally, because vocal communication between dogs and humans is not unilateral, we should consider how dogs perceive human voices in order to fully appreciate their ability to assess vocal signals. Just as humans can extract information from dog vocalisations, so dogs can extract information from human vocal signals. In the final section of this chapter, we discuss evidence that dogs appear to respond to both verbal (syntax and semantics) and subverbal information (vocal identity and emotional prosody) in human speech, as well as to human non-verbal vocalisations (such as cries and laughter).

5.2 HOW DOGS PRODUCE VOCAL SIGNALS

To explore the communicative function of domestic dog vocalisations, we need to first understand more generally how specific acoustic parameters can encode relevant information about the signaller. An efficient way to do this is to consider the structure of vocalisations in the light of their mechanism of production. In most mammals, including dogs and humans, the anatomy of the vocal apparatus is fundamentally similar, so that theories of vocal production can be applied across different species. The source–filter model of human speech production (Fant, 1970; Titze, 1994) has successfully enabled bioacousticians to interpret the acoustic structure of many vertebrate vocal signals within the context of production, enabling researchers to determine how the structure of signals is influenced by the physical and physiological attributes of the caller. According to this model, the production of vocal signals involves independent contributions from two different parts of the vocal apparatus, the *source* and the *filter*.

The *source* of mammalian vocal production is the larynx, a mostly cartilaginous organ that is situated low in the throat where the oesophagus and trachea

join (see Figure 5-1). At the superior border of the larynx, protected by the elastic cartilage of the epiglottis, the glottis consists of soft tissue layers of muscle and vocal ligament (known as the vocal folds) and the spacing between them. The production of vocalisations begins here: air from the lungs forces its way through the closed glottis and the vocal folds are pushed apart. Biomechanical forces then cause the vocal folds to snap shut again, and this sequence of opening and closing of the glottis causes a cyclic and self-sustaining variation in air pressure across the larynx (see Titze, 1994). The resulting waveform is known as the source signal or glottal wave. The rate of oscillation of the vocal folds determines the fundamental frequency (henceforth F0), and associated harmonics of the source signal, and is perceived as pitch by human listeners. F0 is primarily determined by the length and mass of the vocal folds: longer and heavier vocal folds vibrate at a slower rate than smaller vocal folds. In humans, these properties can, to a certain extent, be manipulated by flexion/relaxation of the muscles controlling the lengthening/shortening and tension of the vocal folds. Other characteristics of the source signal include tempo, duration, and amplitude contour, all of which are controlled by a variety of muscular interactions and changes in airflow or subglottal pressure (Titze, 1994). Source characteristics can thus vary between and within vocalisations from the same caller either on a volitional (intonation in human speech: Ohala, 1984; Banse & Scherer, 1996) or on an involuntarily basis (emotional expression in humans: Ohala, 1996; Aubergé & Cathiard, 2003; affective state in baboons: Rendall, 2003; stress in pigs: Düpjan et al., 2008). The implications of this are discussed in more detail later in the chapter.

The source signal then travels through the vocal tract *filter*, which is composed of the caller's pharyngeal, oral, and nasal cavities, before it is radiated into

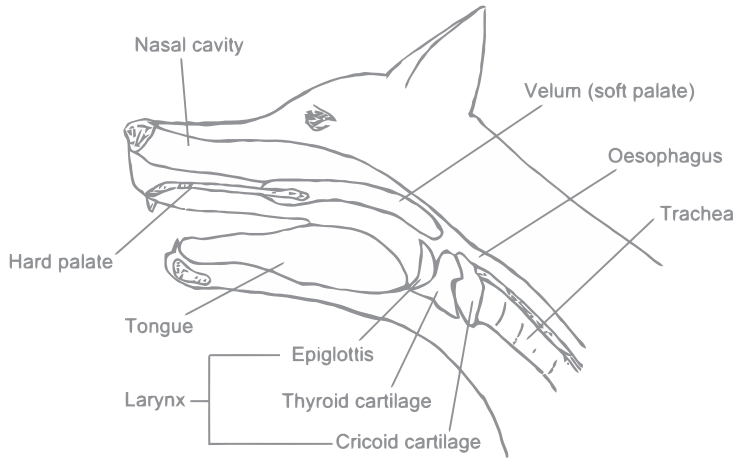

FIGURE 5-1 Major components of the vocal apparatus of the domestic dog. Note that only the outer structures of the larynx are represented in this diagram. *(Based on work by Piérard, 1963.)*

the environment through the mouth and/or nostrils. The vocal tract is referred to as the 'filter' because its resonant properties (determined primarily by its length and shape) selectively enhance or dampen some of the harmonic frequencies of the source signal. Those frequencies that resonate well form spectral peaks, or formants (Fant, 1970). Their perceptual quality is best described as the 'depth' or 'timbre' of a vocalisation, which is not to be confused with its pitch (Titze, 1994). The degree to which formant positioning can be actively controlled is dependent on the species. Humans, for example, demonstrate highly sophisticated neuromotor control of their vocal apparatus: subtle changes in the position of the pharynx, velum (soft palate), tongue, and lips influence the resonant properties of the vocal tract, affecting the relative frequency position of formants, which is necessary for speech production (see Lieberman & Blumstein, 1988; Titze, 2000). In contrast, most mammals, including dogs, do not appear to have fine motor control over their vocal tract characteristics and so the resonant properties of the vocal tract are usually more predictable. Formant frequencies in most non-human calls therefore appear as evenly distributed horizontal bands, often directly reflecting the length of the vocal tract (Fitch, 1997). This distribution of formants is quantified under the term 'formant dispersion' (Fitch, 1997; Reby & McComb, 2003) and constitutes one of the key factors affecting the perceived quality of the vocalisation.

The source- and filter-related acoustic components in mammalian vocalisations are therefore constrained by different production characteristics. By applying this framework to the dog's vocal repertoire, we can explore how dogs produce the acoustic features which characterise their different call types. Moreover, we can see how this influences the functional content of different calls and discuss their evolutionary origins (see Taylor & Reby, 2010, for a detailed review of how the source–filter theory can be applied as a generalised conceptual and methodological framework to investigate vocal communication in non-human mammals).

5.3 DESCRIPTION OF DOG VOCAL REPERTOIRE (AND COMPARISON WITH WOLVES, *CANIS LUPUS*)

Domestic dogs produce a range of vocalisations, many of which are likely to be familiar to non-expert human listeners. We will now provide an overview of the dog's most common vocal signals, focussing both on production aspects and on their evolutionary functionality. In this regard, it is useful to draw comparisons between dogs and equivalent calls in the ancestral wolf. At a first glance, the acoustic structure of dog vocalisations appears to be very similar to that of wolves; several studies present acoustic analyses of calls from both species, and on a structural level, one would be hard-pressed to distinguish between the barks or growls of dogs and wolves (Feddersen-Petersen, 2000; Schassburger, 1993). However, when we look at the context of emission, it appears that the vocalisations of these species may not be the same on

a functional level. Indeed, it is likely that the domestication of dogs has had an important impact on how their vocalisations are used. A broad comparison between the vocal behaviour of domestic dogs, wolves, and other canids is presented in Table 5-1.

The four most common characteristic calls produced by dogs, namely barks, growls, whimpers, and howls (illustrated in Figure 5-2), appear to be shared across breeds (Cohen & Fox, 1976). While we primarily focus on these four call types, we also briefly discuss anecdotal reports of other vocalisations, including both breed-specific calls and those generally occurring across breeds, and highlight where future investigation would be particularly relevant.

5.3.1 Barks

One of the most stereotypical vocalisations of the domestic dog is the bark. Barks are short, plosive signals that can be produced as part of a sequence or in isolation, with an F0 range that can vary considerably across breeds, individuals, and contexts (Cohen & Fox, 1976; Feddersen-Petersen, 2004). The acoustic output is achieved by the dog lowering the larynx and raising the velum, thereby closing off the nasal passage. The sound is then emitted from the mouth, with an open jaw, as shown in Figure 5-3a. Barks are often described as noisy or chaotic because the harmonic-to-noise ratio varies due to the irregular oscillation of the vocal folds (Riede & Fitch, 1999; also see Figure 5-2a). Harrington and Mech (1978) proposed that wolves bark as an aggressive signal, advertising their willingness to defend themselves and their companions or territory. This instinctive response to draw attention to potential intruders remains common in domestic dogs today and is highly likely to have been one of the factors facilitating the association between early hominids and canids. Yet despite similarities, the barking behaviour of dogs and wolves has clearly undergone different selection pressures. While adult wolves bark in specific aggressive contexts, adult dogs appear to bark readily across many different contexts, including playful and positive situations (Cohen & Fox, 1976). This vocal distinction remains apparent even when individuals of both species have been raised together in similar environments, suggesting that the ontogeny of barks in adult domestic dogs has radically changed from its ancestral, context-specific form (Frank & Frank, 1982; Feddersen-Petersen, 2000).

There also appear to be breed-typical differences in the production of barks. In an analysis of vocalisations from nine different breeds of dog, Feddersen-Petersen (2000) found a high level of variability in barking between breeds, with each breed showing between 2 and 12 subtypes of barking based both on their spectrographic features and on behavioural correlates. This suggests that domestication may have affected the vocal behaviour of different breeds in different ways, potentially due to selection by humans for different behavioural roles. Certainly, the observation that dogs bark in many different situations led earlier researchers to speculate that dog barks were a ‘hypertrophy’ of a previously functional behaviour that

TABLE 5-1 Comparative Overview of Dog and Wolf Common Vocalisations

Signal	Acoustic Features		Context	Other Canids Known to Produce this Call Type
Bark	Plosive Variable pitch Long distance	Domestic dog	Non-specific All contexts	Dingo, *Canis lupus dingo* (Corbett, 1995) New Guinea singing dog (Koler-Matznick et al., 2003) Coyote, *Canis latrans* (Cohen & Fox, 1976) Ethiopian wolf, *Canis simensis* (Sillero-Zubiri & Gottelli, 1994) Golden jackal, *Canis aureus* (Estes, 1991) Side-striped jackal, *Canis adustus* (Estes, 1991) Black-backed jackal, *Canis mesomeias* (Moehlman, 1983; Estes, 1991) Dhole, *Canis alpinus* (Volodin et al., 2001) African wild dog, *Lycaon pictus* (Robbins, 2000) Crab-eating fox, *Cerdocyon thous* (Brady, 1981) Hoary fox, *Lycalopex vetulus* (Sillero-Zubiri et al., 2004) Maned wolf, *Chrysocyon brachyurus* (Brady, 1981) Bush dog, *Speothos venaticus* (Brady, 1981)
		Grey wolf	Specific Aggressive	
Growl	Harsh/ broadband Low pitch Short range	Domestic dog	Non-specific Agonistic Playful	Dingo, *Canis lupus dingo* (Déaux & Clarke, 2013) New Guinea singing dog (Koler-Matznick et al., 2003) Coyote, *Canis latrans* (Cohen & Fox, 1976) Ethiopian wolf, *Canis simensis* (Sillero-Zubiri & Gottelli, 1994) Golden jackal, *Canis aureus* (Estes, 1991) Side-striped jackal, *Canis adustus* (Estes, 1991) African wild dog, *Lycaon pictus* (Robbins, 2000) Short-eared dog, *Atelocynus microti* (Sillero-Zubiri et al., 2004) Crab-eating fox, *Cerdocyon thous* (Brady, 1981) Maned wolf, *Chrysocyon brachyurus* (Brady, 1981) Bush dog, *Speothos venaticus* (Brady, 1981)
		Grey wolf	Specific Agonistic	

Whimper/ Whine	Tonal High pitch Short range	Domestic dog	Specific Social distance reducing	Dingo, *Canis lupus dingo* (Déaux & Clarke, 2013) New Guinea singing dog (Koler-Matznick et al., 2003) Coyote, *Canis latrans* (Cohen & Fox, 1976) Ethiopian wolf, *Canis simensis* (Sillero-Zubiri & Gottelli, 1994) Golden jackal, *Canis aureus* (Estes, 1991) Side-striped jackal, *Canis adustus* (Estes, 1991) Black-backed jackal, *Canis mesomeias* (Moehlman, 1983; Estes, 1991) Dhole, *Canis alpinus* (Volodin et al., 2001) African wild dog, *Lycaon pictus* (Robbins, 2000) Crab-eating fox, *Cerdocyon thous* (Brady, 1981) Maned wolf, *Chrysocyon brachyurus* (Brady, 1981) Bush dog, *Speothos venaticus* (Brady, 1981)
		Grey wolf	Specific Social distance reducing	
Howl	Harmonic Frequency modulated Long distance	Domestic dog	Non-specific Function unclear, possibly vestigial response to high-pitched sounds and social isolation	Dingo, *Canis lupus dingo* (Corbett, 1995) New Guinea singing dog (Koler-Matznick et al., 2003) Coyote, *Canis latrans* (Cohen & Fox, 1976) Ethiopian wolf, *Canis simensis* (Sillero-Zubiri & Gottelli, 1994) Golden jackal, *Canis aureus* (Estes, 1991; Jaeger et al., 1996) Side-striped jackal, *Canis adustus* (Estes, 1991) Black-backed jackal, *Canis mesomeias* (Moehlman, 1983; Estes, 1991) Dhole, *Canis alpinus* (Volodin et al., 2001) Crab-eating fox, *Cerdocyon thous* (Brady, 1981) Maned wolf, *Chrysocyon brachyurus* (Brady, 1981)
		Grey wolf	Specific Social cohesion	

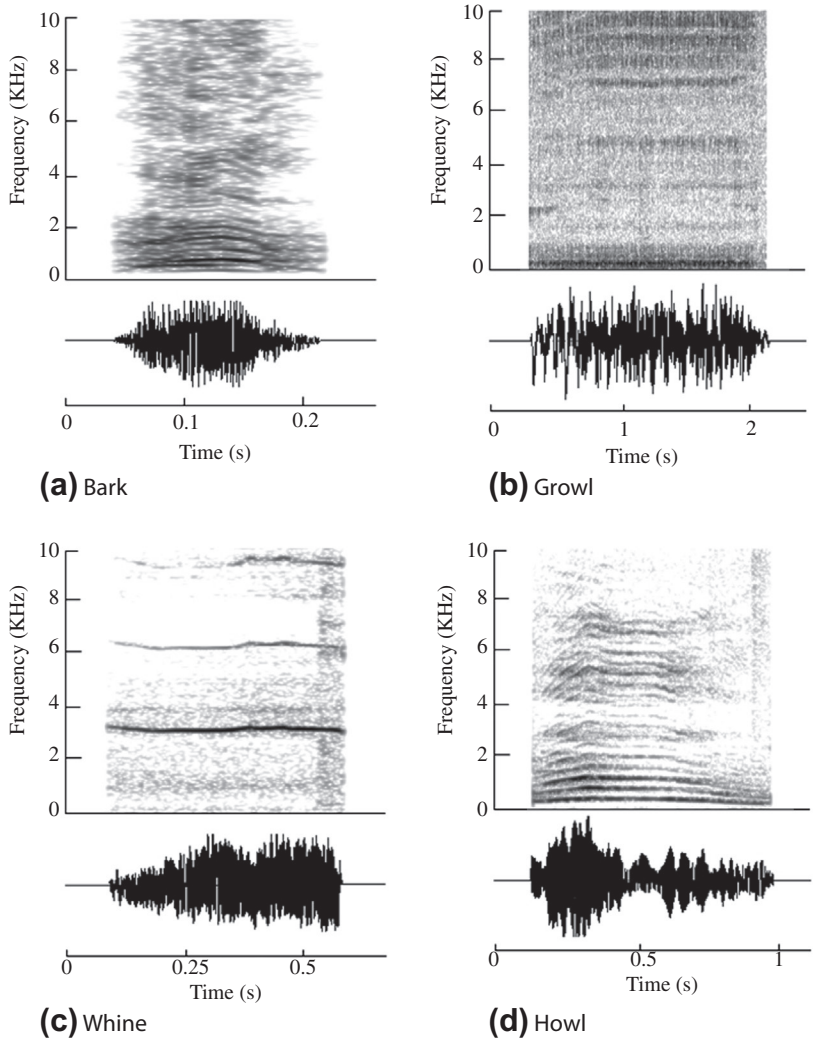

FIGURE 5-2 Spectrograms of a bark (a), growl (b), whine (c), and howl (d).

may have developed due to relaxed selection pressures (Cohen & Fox, 1976). For a long time, it was believed that dog barks were a functionally insignificant, 'non-communicative' by-product of domestication (Coppinger & Feinstein, 1991; Bradshaw & Nott, 1995). However, recent studies have demonstrated that barks do, in fact, broadcast reliable information to receivers, which we discuss in more detail in the following sections.

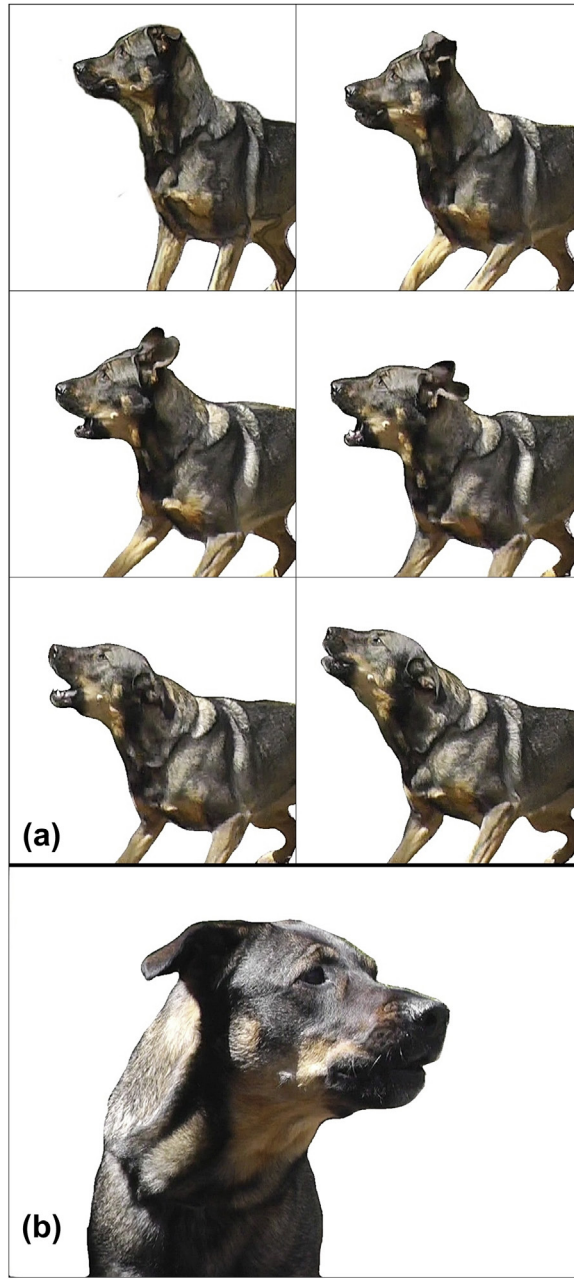

FIGURE 5-3 Tess, a crossbreed, emitting a single bark in response to an unfamiliar person (a) and a growl in response to an unfamiliar dog (b).

5.3.2 Growls

The next most common vocalisation, likely to be just as familiar to most human listeners, is the growl. The growls of dogs and wolves are structurally identical: harsh, low-frequency broadband vocalisations, with an F0 range between 80 and 300 Hz (Cohen & Fox, 1976; Harrington & Mech, 1978; see Figure 5-2b). In contrast to the plosive barks that can propagate over longer distances, it seems likely that growls evolved for the purpose of close-range communication (Cohen & Fox, 1976). In adult wolves, growls commonly occur during confrontations and dominance interactions and near the site of a fresh kill (Fox, 1984), and they may also be used to warn off subordinate individuals or scavengers. Growls accompanied by threatening body posture are more likely to be followed by an actual attack both on other wolves and on humans (Rutter & Pimlott, 1968; Harrington & Mech, 1978). In subadult wolves, growls are also commonly emitted during play fights. Fox (1984) observed that dominant cubs play growled more than subordinate cubs, whereas subordinate cubs whimpered more, indicating that play growling could be part of the maturation process as juvenile wolves practise behaviours that will be needed in adulthood. In contrast, play growls are not heard during interactions between adult wolves (Fox, 1984). This represents a clear distinction between the vocal repertoires of wolves and dogs. Indeed, while adult dogs clearly growl in aggressive contexts (Figure 5-3b shows Tess, a crossbreed, emitting an aggressive growl with closed jaw and drawn-back flews in response to another dog), they also frequently growl in playful situations, mirroring behaviour observed in juvenile wolves (Cohen & Fox, 1976). In adult dogs such play growling can occur during positive interactions with humans, dogs, other animals, or inanimate objects (e.g., toys), and so here too it seems likely that the relaxation of natural selection triggered by domestication has affected the function of the vocalisation.

5.3.3 Whimpers and Whines

The terms ‘whimper’ and ‘whine’ are taken here to refer to high-pitched, tonal vocalisations that can be emitted through the nasal cavity (Fitch, 2000b; Figure 5-2c). In both dogs and wolves, they are produced in a greeting context and as a solicitation for food or care (Fox, 1984; Cohen & Fox, 1976). For the production of whimpers, the larynx remains inserted in the nasopharynx, and the velum position remains open (Fitch, 2000b). In wolves, greeting whimpers might be exchanged between adults or between cubs and their dam. In dogs, similar greeting whimpers are mostly likely to occur when a familiar human is approaching (Cohen & Fox, 1976). It is as yet unknown whether this mirrors cub–dam behaviour in the same way as face licking and play bowing (Bekoff, 1974), or whether it is a greeting from a subordinate adult to a dominant individual. Importantly, whimpers and whines are produced exclusively by a non-aggressive caller and often serve to appease a dominant individual and decrease

social distance. Based on this observation, it can be proposed that whimpers serve as the functional opposites of growls.

5.3.4 Howls

Howls are frequency-modulated harmonic vocalisations that are perhaps more commonly associated with wolves than with domestic dogs. Indeed, wolves are well known for their group howling, a harmonious chorus containing both howls and barks and able to propagate across wide distances up to two kilometres (Harrington & Mech, 1978). Recent vocal analyses in Italian wolves found that the mean F0 of howls emitted by the same pack could range from 274 to 908 Hz (Zaccaroni et al., 2012) and there is some speculation that the frequency range broadcast by a chorus advertises the presence and location of a pack, thereby enabling different packs to minimise the risk of meeting each other (Joslin, 1967). In addition to chorus howling for territory maintenance, wolf howls may enable separated individuals to regroup, possibly because they are individually recognisable by other group members (Harrington & Mech, 1978; Theberge & Falls, 1967). A recent study found that captive wolf pack howling rates were higher when dominant individuals or close social partners were absent from the group. Whilst the remaining pack members' stress hormone (cortisol) levels were higher when the dominant individuals were missing, there was no difference in cortisol levels when preferred pack mates were absent. This suggests that howl production is not entirely driven by the physiological or emotional state of the signaller, but rather is to some extent under voluntary cognitive control and can be flexibly emitted depending on the social context (Mazzini et al., 2013). Coyotes, jackals, and domestic dogs also howl (see Figure 5-2d), although the informational content and function of howls in these species and specifically domestic dogs remains largely unstudied.

5.3.5 Other Vocalisations

In addition to the four main call types described in the preceding sections, dogs produce more unusual, sometimes breed-specific vocalisations. The work of Feddersen-Petersen (2000) and others suggests that domestication may have affected the vocal behaviour of different breeds in different ways. Indeed, selective breeding for specific behavioural roles appears to have involved the selection of vocal behaviour. In some breed groups, this may have influenced the rate of occurrence of shared call types. For example, guarding breeds were selected for protective instinct and thus display strong defensive behaviours, including exaggerated barking and growling. Terriers were likewise bred to bark frequently in order to alert hunters to their location whilst pursuing prey animals in underground burrows, but gundogs must remain silent whilst hunting so as not to frighten away any game (Coren, 2005). Moreover, distinct vocalisations may have developed in specific breeds. Scent hounds, for example, appear to

remain in constant communication with each other while hunting using a very distinct 'baying' vocalisation, a melodious call containing many tonal variations (Coren, 2005). The propensity to produce unusual call types appears to be higher in breeds descended from the Indian plains wolf (*Canis lupus pallipes*) such as the dingo, New Guinea singing dog, and basenji dog, which all show a number of unique attributes that do not occur in other breeds. Examination of the vocal tract of the New Guinea singing dog revealed a rudimentary two-lobed uvula (a mass of tissue that is suspended from the velum and hangs above the throat), which has not been observed in any other canid. It is thought that the vibration of the uvula may allow these dogs to produce their distinctive 'trill', a high-frequency pulsed vocalisation (Koler-Matznick et al., 2003). Basenjis, an African breed selected for pack hunting large game, are known to produce a high-pitched, harmonic call known as a 'yodel' or 'baroo'. A study of two basenji dogs revealed that their laryngeal anatomy was significantly different from that of other breeds (Ashdown & Lea, 1979). Similarly, the dingo also produces a distinctive stereotyped composite 'bark-howl' (Deaux & Clarke, 2013). More research is needed to fully understand these breed-specific vocalisations.

Overall, despite some functional differences, it is clear that the vocalisations of domestic dogs and wolves follow a broad, predictable pattern of acoustic output, with lower-pitched sounds occurring more commonly in aggressive contexts and higher-pitched sounds occurring in social and playful contexts. In the previous section, we noted that the relaxed selection pressures of domestication may have contributed to a change in the context-specificity in which barks and growls are produced. On a structural level, these vocalisations are indicative of agonistic displays, yet they are heard across many different contexts. In the next section, we review investigations of the acoustic structure of vocal signals aimed at understanding their information content. We focus primarily on barks and growls because these two call types are the most well-documented and researched vocalisations of the dog.

5.4 THE COMMUNICATIVE FUNCTION OF ACOUSTIC VARIATION

5.4.1 Indexical Information

Within the methodological framework of the source–filter model, researchers interested in mammal communication can explain why different calls contain particular acoustic features. Any vocalisation in which the acoustic properties are predictable as a function of the physical attributes of the caller has the potential to broadcast accurate information about that physical attribute (Fitch, 1997; Fitch & Reby, 2001). It is well documented that larger-sized individuals are at an advantage over smaller individuals during agonistic encounters, and individuals benefit from avoiding costly fights that they are unlikely to win (Schmidt-Nielsen, 1975; Peters, 1986). Acoustic advertising and assessment of body size

are thereby central to social interactions in some species (e.g., red deer: Reby & McComb, 2003), and identifying the relevant acoustic correlates to body size is thus an important task in vocal communication research.

Because of the perceptual salience of pitch to human listeners (Ohala, 1984), some earlier studies of mammal vocal communication focussed on F0 as a potential source of size-related information (see August & Anderson, 1987). As we saw earlier, F0 is primarily dependent on the length and mass of the vocal folds. Generally speaking, the soft tissue anatomy of the vocal folds means that they are not stringently constrained by the body size of the individual (Fitch, 2000c), but it has been observed that across age and sex categories (possibly due to age-related vocal-fold growth and sexual dimorphism) F0 can sometimes be broadly correlated with caller body size. For example, in both baboons and humans, males are larger than females and have a lower F0 (Rendall et al., 2005; Puts et al., 2006). In relation to other mammals, canid species have a comparably low level of sexual dimorphism, which is related to their largely monogamous breeding systems (Bekoff et al., 1981). The grey wolf is the most sexually dimorphic canid species; males have around 18% greater body mass than females (Hillis & Mallory, 1996). Within dog breeds, those of comparable size to wolves show similar levels of body mass dimorphism, which was found to relate to a difference of 5%–7% in vocal tract length (VTL) between adult males and females for Portuguese water dogs (Plotsky et al., 2013). However, the size difference between sexes decreases in smaller breeds (Frynta et al., 2012), and no sex-related differences in vocal anatomy or in F0 have been identified across breeds when controlling for body weight (Riede & Fitch, 1999; Taylor et al., 2008). However, artificial selection by humans has led to an exceptionally large level of size variation between breeds, giving dogs the highest level of morphological variation of any mammal (Wayne, 1993). Consequently, across breeds, F0 can provide a broad indication of body size, with smaller breeds producing growls with a higher F0 than larger breeds (Taylor et al., 2008). On an individual level, F0 is influenced by changes to respiration and muscular control of the tension of the vocal folds. Therefore, F0 is highly variable both within and between calls, across individuals, and across breeds (Feddersen-Petersen, 2000; Yin, 2002; Taylor et al., 2008).

In contrast, the vocal tract cannot grow independently of the rest of the body, as its development is anatomically constrained by skeletal structures (Fitch, 2000a, 2000c). Vocal tract length is thus more directly dependent on body size, and investigations have confirmed a strong positive correlation between anatomical VTL and body size in many species, including domestic dogs (Riede & Fitch, 1999). As formant dispersion is controlled by VTL, this means that unlike F0, formants have the potential to provide accurate or 'honest' information about the caller (Fitch, 1997; Reby & McComb, 2003). A direct negative correlation between formant dispersion and body size has been confirmed in the signals of many species, including dog growls (Riede & Fitch, 1999; Taylor et al., 2008). When we analyse formant dispersion within any growl, it

is thus possible to predict the body size of the caller. It is also noteworthy that amongst dog vocalisations, the growl is particularly suited to transmitting size-related information. The harsh, broad bandwidth structure of growls transmits acoustic energy across a wide range of frequencies, which increases the perceptual salience of the formants (also see Ryalls & Lieberman, 1982). As growls evolved, this aspect of their acoustic structure may have originally been selected for in the context of agonistic interactions because it most effectively advertised body size to competitors.

In terms of perceptual relevance, several experiments show that domestic dogs are sensitive to size-related acoustic variation in conspecific vocalisations. Taylor et al. (2010a) observed that dogs responded to variations in the formant dispersion of growls in a manner consistent with size assessment. In the absence of visual size cues, dogs' behavioural responses to playbacks simulating small dogs were significantly different than their responses to playbacks simulating large dogs. Responses were additionally influenced by the subjects' own body size. Indeed, large dogs showed more motivation to investigate when growls simulating a smaller intruder were played than when growls simulating a larger intruder were played. In contrast, small dogs did not respond differentially to small versus large intruders and reacted in a more consistent manner to all playbacks (Taylor et al., 2010a). Taylor and colleagues' observations of a potential effect of the relative size of signallers and receivers are consistent with the results of an earlier study of visual communication in dogs (Leaver & Reimchen, 2008). These authors found that only dogs larger than a Labrador retriever model responded differentially depending on modified tail positions in the latter, while dogs smaller than the model did not respond differently depending on tail position and, in fact, showed lower motivation to approach or interact with the model overall. The two studies strongly suggest that such responses may not be universal across dogs but may be influenced by their own body size. In fact, small dogs (under 10 kg) are the most likely victims of bites and injuries inflicted by other dogs (Shamir et al., 2002) and may therefore benefit the most from avoiding interactions with unfamiliar individuals. Because domestication is likely to have eliminated most competition for resources, we can speculate that small dogs may have little benefit to gain from a confrontation with a potentially aggressive intruder. As such, assessing caller body size on the basis of auditory information may well be a functionally relevant ability even within the context of a domestic animal.

Integrating auditory information with related visual information is an ability most humans take for granted, and there is increasing evidence that many species are able to process some types of information across modalities (Adachi & Fujita, 2007; Adachi et al., 2007; Ghazanfar et al., 2007; Proops et al., 2009). In two separate experiments using cross-modal designs, dogs demonstrated that they were able to reliably match acoustic size information in growls with corresponding visual information [for the visual stimuli, Faragó et al. (2010) used images of a small or large variant of a dog projected onto a white screen,

whereas Taylor et al. (2011) used taxidermies of a Jack Russell terrier and a German shepherd dog]. In both studies, dogs preferentially looked towards the visual stimulus that matched the apparent acoustic size of the caller, showing not only that they had the ability to perform this association, but also that they were motivated to do so, suggesting that this task was functionally relevant to dogs.

Size-related acoustic variation may moreover be functionally relevant across species. In a psychoacoustic investigation, Taylor et al. (2008) demonstrated that human listeners are able to accurately attribute associated body size to growls that had been resynthesised to vary only in formant dispersion, with all other acoustic parameters standardised across stimuli. Interspecific size assessment is also interesting to consider in the context of the predatory potential of domestic dogs. As larger predators are more successful hunters and generally outperform smaller predators (Gittleman, 1989; Herrel & Gibb, 2006; MacNulty et al., 2009), the body size of predators is valuable information to any prey animal. In support of this, species at risk of predation by dogs such as domestic sheep or red deer appeared to respond to playbacks of domestic dog growls in a manner consistent with size assessment. In a playback experiment, the individual vigilance of groups of domestic sheep and individual red deer hinds was found to be greater in response to growls typical of large dogs than to growls typical of small dogs (Taylor & Reby, unpublished data). A confirmation of these observations would bring further support to the theory that size perception in the auditory domain is universal across mammalian species and applicable in a number of different functional contexts.

Attending to information about body size and using it to assess the physical attributes of potential competitors can thus have important implications for the reproductive opportunities and survival of receivers. As such, it is very likely that the functional decisions of receivers, based on this information, may place additional selection pressure on signals at the level of their production. In an influential paper based on a comparative study of vocalisations used in agonistic displays in a range of mammalian and avian species, Morton (1977) suggested that audible frequency differences in vocalisations reflect ritualised signalling: animals with aggressive motivation produce low-pitched, broadband vocalisations (such as growls and hisses), whereas animals with affiliative or submissive motivation produce high-pitched vocalisations (such as whimpers and whines). This theory, known as Morton's motivation-structural code, is based on the observation across several species that aggressive and dominant animals seek to project (both visually and acoustically) a larger impression of body size and greater threat, whereas friendly or submissive animals seek to project a smaller impression of body size and lesser threat (Morton, 1977; Ohala, 1984; Owings & Morton, 1998).

Acoustic variation of signals encoding information that is associated with attributes such as body size may thereby become ritualised, broadcasting more transient information pertaining to the motivational state of the caller (Ohala,

1984; Taylor & Reby, 2010). While mechanisms for visual maximisation or exaggeration of body size such as piloerection is relatively common among mammals, few acoustic equivalents have been reported (although see the retraction of the larynx in red deer; Reby & McComb, 2003). The use of lower frequency vocalisations could thus be representative of aggressive motivation not just across call types as described by Morton (1977) but within call types as a ritualised function of the acoustic size code (Ohala, 1984, 1996; Taylor & Reby, 2010).

5.4.2 Dynamic Information

Whilst dogs do not appear to modify the formant frequency scaling in their vocalisations, and thus only seem to provide a static indication of body size, they are able to produce an impressive range of F0 both within and between calls (Yin, 2002). Variations in the source signal may provide reliable information about the affective state of callers because it is influenced by changes to rate of respiration and/or muscular tension in the vocal folds (Hauser, 2000; Bachorowski & Owren, 2008), both of which are linked to physiological arousal. One outcome of this is that these temporal characteristics can be affected by state of arousal. Indeed, in barks and growls, inter-call interval and call duration are affected by production context: barks recorded in an aggressive context (being approached by a stranger) were found to have very short inter-bark intervals compared to barks recorded in a playful context (Yin, 2002), whereas growls recorded in an aggressive context are characterised by a low calling rate and have long call durations (Taylor et al., 2009). Temporal characteristics might thus provide a guide to motivational state, although the salience of this to listeners is yet to be established. Another, better understood, outcome is the effect of arousal on F0 in the context of broadcasting of motivational state information. Domestic dog barks occur on a graded frequency scale: lower peak, mean, and range of F0 are found in barks recorded in an aggressive situation than in those recorded in a playful situation (Feddersen-Petersen, 2000; Yin, 2002). This has been confirmed in single-breed (Pongrácz et al., 2006) and mixed-breed experiments (Taylor et al., 2009) and is sufficiently predictive that barks can be reliably categorised as aggressive or playful on the basis of statistical analysis (Yin & McCowan, 2004). Moreover, Pongrácz et al. (2005) showed that human listeners are able to discriminate between barks recorded when dogs were approached by a stranger, antagonised by a Schutzhund trainer, going on a walk, left in isolation, or playing. In addition, listeners attributed emotional content to barks from the preceding contexts (aggressiveness, fear, playfulness, or happiness; Pongrácz et al., 2006). Similarly, dogs appear to physiologically discriminate between barks from unfamiliar dogs recorded in two different situations, but the functional relevance of their discrimination is unclear because dogs did not respond with quantifiable behaviours (Marós et al., 2008).

Growls have similarly been studied within the framework of the acoustic size code hypothesis. Investigating the theory that more aggressive dogs may manipulate their formants to sound larger, Taylor et al. (2009) did not find any difference in the formant dispersion of growls recorded in an aggressive (approaching human) or playful context. However, Faragó et al. (2010) analysed growls recorded in an extremely high emotional valence situation (food guarding against an unfamiliar dog) and found that formant dispersion in these growls was, in fact, representative of size maximisation. A cineradiographic study showed that domestic dogs may be able to do this by modifying the level of nasalisation of a signal via small jaw and tongue movements (Fitch, 2000c). In terms of perceptual significance, dogs avoided food when they were presented with food-guarding growls but did not avoid food when presented with either playful growls or defensive (threatening stranger approach) growls (Faragó et al., 2010). It is noteworthy that manipulating size-related formant frequency scaling can also potentially affect the perception of motivation by humans, despite the lack of relationship between formant and motivation in dog growls. Indeed, humans attributed higher levels of aggressiveness to growls where formants indicate larger dogs than to growls where they indicate smaller dogs (Taylor et al., 2010b). This suggests that there may be a pre-existing perceptual bias in humans to confound information about body size and motivation in the acoustic domain (possibly such a bias may have developed because of the human tendency to use size-related acoustic variation in voice to rate traits such as masculinity and dominance; e.g., Puts et al., 2007) and that this bias occurs despite, or perhaps in addition to, their ability to accurately judge caller body size (Taylor et al., 2008, 2010b).

Overall, the acoustic variation in barks appears to be better suited to broadcasting information about dynamic state than the acoustic variation in growls. However, size-related information in growls may in some cases of very high emotional valence serve to broadcast a highly aggressive motivation. Similarities in signal production across different individuals allow receivers to make reliable judgements about the dynamic and static attributes of the signaller; however, differences between individuals in vocal anatomy and physiology mean that some level of individual variation within calls still occurs. If these differences are both static and perceived by receivers, it becomes possible to identify individuals from their calls.

5.4.3 Individual Recognition

Some mammalian vocalisations may contain a vocal signature, i.e., a unique combination of acoustic parameters that enables receivers to recognise or identify the signaller. In an early experiment, Yin and McCowan (2004) found that dog barks could be statistically attributed to individual callers (irrespective of production context) using a discriminant function analysis. Similarly, Molnár et al. (2008) developed computer-based learning software, which categorised

individual dogs on the basis of their barks. Human listeners, however, struggled to reliability discriminate between barks from dogs of the same breed (Molnár et al., 2006), although the presentation of five-bark sequences did somewhat improve discrimination. Molnár et al. (2009) suggest that the interaction between harmonic-to-noise ratio and mean F0 may be individually different between dogs and that although these cues may be too subtle for human listeners, they may be perceptually accessible to other dogs. To date, however, no study has conclusively demonstrated individual recognition on the basis of auditory information in dogs.

Although it is possible to ask human listeners what information they can extract from dog vocalisations, innovative paradigms are needed to determine how this information is both assessed and used by other dogs. Whilst this leaves many areas for advancement in determining how dogs perceive information within their own species' calls, it is possible to look further into how dogs may have adapted these perceptual and cognitive abilities to extract information from human vocal signals.

5.5 THE PERCEPTION OF HUMAN VOCAL SIGNALS BY DOGS

So far, it is evident that dog vocalisations can broadcast relevant information about the physical and motivational state of signallers, and that listeners may make functionally useful assessments on the basis of this information. However, domestication has strongly exposed dogs to the human environment, and thus, it is interesting to consider what perceptually relevant information dogs may be able to extract from human vocal signals. It has long been known that domestic dogs are highly sensitive to the visual cues used in human social communication, learning to use pointing gestures and gaze direction to locate hidden rewards from an early age (e.g., Miklósi et al., 1998; Hare & Tomasello, 1999). Furthermore, when asked about their dogs' responsiveness to verbal utterances, owners tend to state that their dogs are good at interpreting communicative intentions in speech. This belief often leads owners to attempt to communicate verbally with their dogs, but what information do dogs actually perceive when they are spoken to?

Although the human voice contains the same anatomically related acoustic components as other mammals, providing indexical and dynamic information about the speaker, the most crucial dimension in human vocal communication is the ability to transmit semantic information through speech. Speech consists of hierarchical segments, starting at the basic level of consonants and vowels, or phonemes. These are produced through intentional movements of the vocal apparatus, which alter the relative positions of the lower formant frequencies (Fant, 1970). Phonemes are then combined to create meaningful words and structured into sentences, allowing humans to communicate complex language in a spoken format. Human vocalisations thus have the potential to transmit several levels of information: indexical, emotional, and semantic. We next discuss

the ability of domestic dogs to extract both the shared and unique levels of information from the human voice.

5.5.1 Indexical Information

The anatomical features of the human vocal apparatus differ in size between children and adults and between adult men and women (Fitch & Giedd, 1999), generating categorical acoustic differences in F0 and formant dispersion that provide an indication of the age and sex of the speaker (Smith & Patterson, 2005). Because dogs perceive indexical information broadcast in the calls of their own species, they may also have the ability to discriminate the systematic variation between human voices derived from these static attributes. F0- and formant-related gender differences allow human listeners to readily classify adult human voices as male or female (e.g., Coleman, 1976). With discrimination training, dogs can learn to distinguish between the average F0 values of adult male and female voices and to discriminate formant-related variation in the human voice (Baru, 1975). If dogs have sufficient social experience with men and women, they spontaneously learn to categorise human voices as male or female, enabling them to associate unfamiliar voices with people of the same gender (Ratcliffe et al., submitted). This shows that dogs naturally learn to discriminate acoustic variation that relates to the biological attributes of human speakers and use this information to guide their visual search towards a matching person. Further research is needed to determine which acoustic characteristics dogs use to categorise human voices as male or female, and if they also categorise other indexical cues present in the human voice, such as age or body size, to associate voices with different speakers.

5.5.2 Individual Recognition

The co-variation of characteristics that are unique to the vocal apparatus of each individual speaker contributes to the existence of individually distinct voices in humans (Bachorowski & Owren, 1999). Humans are able to match voices with other unique traits of known individuals such as their facial features from infancy (Bahrick et al., 2005), making individual people highly recognisable across sensory modalities. The cognitive mechanisms of individual recognition are less well understood in animals (see Proops et al., 2009), but it has been shown that domestic dogs are able to match auditory and visual cues to human identity, indicating some level of individual recognition. Adachi et al. (2007) illustrated this by using an expectancy violation (congruent/incongruent) paradigm. In this experiment, dogs were first presented with a voice recorded from either their owner or a stranger and were then presented with an image of either their owner's face or a stranger's face. The dogs looked longer at the person's face when it did not match the preceding voice, suggesting that the dogs expected to see their owner when they heard his/her voice, and did not

expect to see their owner if they heard a stranger's voice. It therefore appears that dogs are able to perceive and categorise the individually distinctive acoustic cues characterising their owner's voice and use this information to recognise individual human vocal signatures. In this study, the acoustic stimulus was the dog's own name, a highly familiar signal. To establish the extent to which this ability is independent from learnt signals, we therefore need to determine if dogs can also recognise their owner's voice when he/she is saying unfamiliar phrases. It is also unknown whether dogs are able to discriminate between two familiar people, which would provide further evidence in support of auditory individual recognition.

5.5.3 Emotional Information

In addition to anatomically constrained information, dynamically controlled prosody is also present in human-voiced signals. At an acoustic level, vocal prosody is produced by variation in F0, duration, and intensity, which in addition to clarifying the speaker's intentions (e.g., questions versus statements) can also provide information about his/her emotional state. Different emotions can be related to predictable variation in the relative position and modulation of F0, voice quality, and formant values across speakers, and are universally identified by human listeners (Sauter et al., 2009). It appears that dogs can also discriminate and perhaps even assess prosodic variation generated by different emotions. In one study, dogs were found to respond in a similar manner to vocal commands when spoken in a neutral or positive tone, but their responses became more variable when commands were spoken in a negative tone (Mills et al., 2005). This suggests that dogs can perceive and respond to changes in vocal prosody generated by the emotional state of the speaker. In line with this, Scheider et al. (2011) found that dogs searched longer for food when an experimenter spoke to them in a high-pitched, encouraging voice. However, when spoken to in a low-pitched, imperative voice the dogs reduced their movements and responded more often by sitting or lying down.

One possible explanation of these observations is that dogs learn to associate specific prosodic cues with different behavioural responses, as the dogs produced responses associated with obedience training after hearing prosodic cues that were likely to mimic those used to give previously learnt commands. However, an intriguing alternative is that similarities between the acoustic characteristics in the vocalisations of dogs and humans in specific contexts may facilitate the generalisation of responses across species, in this case causing the dogs to respond submissively to a more negatively valenced voice. Potential support for this explanation can be derived from findings that prosodic cues indicating anger and happiness in human speech may also be related to the acoustic size code discussed earlier in this chapter. Human listeners perceive synthetic vowels created with a dynamically lower F0 and smaller formant dispersion as being spoken in an angry voice, whilst vowels with a dynamically

higher F0 and a larger formant dispersion are perceived as being spoken in a happy voice (Chuenwattanapranithi et al., 2008). This potential universality in ritualisation across mammal vocalisations may allow dogs to generalise their responses to specific prosodic cues, aiding their perception of certain emotions in the human voice. Rather than learning to associate specific responses with different emotional cues in the voice, similarities in the way that emotions are advertised across mammal vocal signals may therefore cause affect–induction (Owren & Rendall, 1997, 2001), influencing the emotional state of the dog and creating an innate empathic response.

Both dog owners and non-dog owners generally agree that dogs are empathically sensitive to human emotions (Vitulli, 2006), and there is some empirical evidence to suggest that dogs may recognise the valence of human emotions. In a preferential choice paradigm in which only one box out of two contained a reward, dogs were more likely to choose a box to which they had seen a person respond with a happy reaction than one paired with a disgusted reaction, demonstrating their ability to use human emotional expressions to guide their choice (Buttlemann & Tomasello, 2012). Interestingly, dogs are better at making such decisions based on emotional expressions of familiar rather than unfamiliar people (Merola et al., 2013). This study also found that dogs' responses are more strongly guided by happy expressions than by avoiding boxes paired with neutral or fearful expressions. This suggests that dogs may learn to associate positive expressions in familiar people with positive outcomes. Finally, it has also been observed that dogs are more likely to approach a person in a submissive manner when that person is pretending to cry than when he/she is talking or humming (Custance & Mayer, 2012), although it is difficult to determine conclusively why dogs may respond differently specifically to crying.

Whilst it remains unclear exactly how dogs process emotion in human voices, they do perceive and appear to make some form of assessment on the basis of the vocal prosody related to human emotional expression. The ability of dogs to recognise prosodic cues in the voice is also likely to be facilitated by the way many dog owners speak to their pets. Just as it is possible that dogs' vocal repertoire has adapted to facilitate interspecific communication with humans (Pongrácz et al., 2005), so humans also change their speech patterns when talking to dogs and other pets. Pet-Directed Speech (PDS) mimics the way in which parents speak to their young children (Hirsh-Pasek & Treiman, 1982). Both Infant-Directed Speech (IDS) and PDS contain a higher mean F0 and a larger F0 range than Adult-Directed Speech (ADS). The higher mean F0 in IDS appears to be an important prosodic feature for engaging an infant's attention and thus promoting and facilitating social interaction (Fernald & Kuhl, 1987). Perhaps, then, it is not surprising that dogs seem to find higher-pitched voices more encouraging in a food-seeking task (Scheider et al., 2011), although to date it has not been determined whether PDS is more effective than ADS in engaging a dog's attention.

Exaggerated prosodic variation also enables adults to recognise emotion more easily in IDS independently of language or culture (Bryant & Barrett, 2007). Although the emotional salience in PDS is not rated as strongly as in IDS, it is significantly higher than in ADS (Burnham et al., 2002). The increased emotional expression in IDS is thought to communicate intentions to the infant (Fernald, 1989), and it may be that the exaggerated emotional expression in PDS also influences responses in dogs. There is some evidence to support this: rising tones are more effective than descending tones in obtaining responses from dogs that require increased movement (McConnell, 1990). Likewise, in IDS, rising F0 contours are associated with gaining attention and encouraging a response (Fernald, 1989). Instead of being a by-product of our parental attachment to dogs, PDS may thus actually share some of the functions of IDS, such as maintaining dogs' attention and manipulating their behavioural responses. Further work is now needed to assess how different intentions are expressed in PDS and whether dogs respond to these cues.

Aside from the possible functions of PDS, the tendency of dog owners to change their speech patterns when speaking to dogs in a way that mimics speech patterns directed towards infants should probably not be unexpected; many dog owners form strong attachment to their dogs, often viewing them as family members (Cain, 1985). In fact, there is some evidence that dogs may also have been selected to have childlike facial features because they trigger parental attachment in their owners (Archer & Monton, 2011). These influences could have an additive effect, causing many owners to attempt to communicate with their dogs as though they were human infants. In support of this, dog owners have been shown to spontaneously produce PDS when interacting with their dogs during situations designed to assess attachment, namely adaptations of Ainsworth's Strange Situation Test, which stimulate caregiving and protective responses. Although both men and women used PDS, the female owners spent significantly more time speaking to their dogs and were more likely to use PDS (Prato-Previde et al., 2006). It is interesting to note, however, that owners may be inherently sensitive to the limited linguistic abilities of their dogs: vowel hyperarticulation, a critical component of IDS thought to be specific to the teaching of language, is absent from PDS (Burnham et al., 2002).

5.5.4 Semantic Information

While, as we just discussed, human voices are comparable with the vocalisations of other mammals in terms of broadcasting indexical and motivational information, humans alone appear to have evolved a further dimension of vocal communication: speech (see Fitch, 2000a). By intentionally manipulating the resonance frequencies within vocal signals, humans are able to produce the phonological structure necessary to create words. These words can be used to refer to abstract concepts, giving them semantic meaning. Furthermore, humans have the ability to generate and process an infinite number of combinations between

words using syntax or grammar, giving speech a rule-based recursive structure. Whilst human speech is therefore very different from canid vocal signals in this respect, in the context of domestication, dogs are likely to benefit from the ability to perceive semantic information in human vocal signals. In order to understand the extent to which dogs are capable of perceiving this information, it is important to review how human speech signals are produced and how semantic information is encoded. To create the precise sounds for different words, humans make specific movements of the articulators, including the tongue and lips, to alter the shape of the vocal tract. This causes dynamic variation in the formant frequencies, particularly those at the lowest frequencies, producing the different phonemes that are categorised as vowels (see Titze, 2000). We have seen that, because dogs have limited ability to control formant positions in their own calls, their formant positions show little variation during vocalisations and are dependent on the length of the vocal tract. We have already seen that dogs are perceptually aware of formant-related information, as they are capable of using the formant scaling in growls to judge the body size of other dogs. This leads to the interesting question of whether dogs can perceive more than the static scaling between formants, and if they are capable of perceiving the dynamic formant modulation used to create the different phonetic sounds in human speech. The first demonstration that dogs are indeed able to perceive the relative positions of individual formants involved training dogs to discriminate between individual vowels (Baru, 1975). Indeed, we now know that discriminating between vowel sounds is important for dogs to learn to identify individual words. Fukuzawa et al. (2005) demonstrated that changing the phonemes embedded in previously learnt commands can lead to a significant decline in the number of correct responses, illustrating that specific phonemes are a crucial part of word recognition for dogs. Although further research is needed to better understand the extent to which dogs perceive and categorise phonemes and how their abilities relate to our own perception, they are clearly able to use phonetic information in order to identify a wide variety individual words.

Through training, dogs can learn a very large number of vocal labels relating to specific objects, responses, and events, comparable with the word acquisition abilities of apes, parrots, dolphins, and sea lions (Miles & Harper, 1994). This was first demonstrated in dogs by Kaminski et al. (2004) with a border collie named Rico. Rico could be given a verbal instruction by his owner to retrieve a specific toy (e.g., 'fetch teddy') from a different room containing an array of 10 familiar toys, and successfully responded to the verbal labels of 200 different toys. Rico also showed evidence of learning new object labels on their first presentation, as when asked to fetch a new toy included amongst an array of 7 familiar toys, Rico successfully excluded the familiar toys and retrieved the new toy. He also showed retention of the new label, as after a month without testing the new object was placed with 4 familiar and 4 unfamiliar toys, and Rico still successfully retrieved the new object. Because the acquisition of the new label was so rapid, it provided evidence that rather than just learning to associate

the word with the object (which often takes many repetitions), dogs could be capable of 'fast mapping', showing the ability to learn that a new word may be used to refer to something new in the environment during a single exposure, a skill that was previously thought to be limited to humans.

However, Bloom (2004) questioned whether Rico understood that words were referential in the same way as humans, as the verbal command could have been represented as a single proposition (e.g., 'fetch the sock'). If this was the case, Rico may have merely associated the command with a specific behavioural response (retrieving the object) without understanding that the label for the object was independent from the action of retrieving it. To test this, Pilley and Reid (2011) taught a border collie named Chaser the labels of more than 1,000 objects and associated these objects with different behavioural commands. Chaser responded to novel combinations of behavioural responses and objects (e.g., 'take ball' versus 'paw ball') from the first trial, illustrating that she was able to differentiate the object label from the action directed towards it. This ability was confirmed by Ramos and Ades (2012) using a crossbreed dog named Sophia, who also learnt to respond appropriately to two-item spoken instructions including both object- and action-related words (e.g., 'stick fetch' versus 'stick point'), even when the word order was reversed. Comprehension of two-item sequences has also been shown in other language-trained species [e.g., African grey parrot (*Psittacus erithacus*): Pepperberg, 1981; California sea lion (*Zalophus californianus*): Gisiner & Schusterman, 1992)], and these sequences also appear to be used naturally in some primate signals [e.g., Diana monkeys (*Cercopithecus diana*) perceive combinatory calls produced by Campbell's monkeys (*Cercopithecus campbelli*): Zuberbuhler, 2002].

Whilst these studies demonstrated that dogs, amongst some other mammal species, can learn to respond to combinatory signals, and therefore show understanding that the label is associated with the object and not the behavioural response, Bloom (2004) also argued that dogs may not recognise that words are symbolic, referring to categories and items in the environment. Although both Rico and Chaser had shown evidence of 'fast mapping', which suggested some understanding of words as verbal referents, potential problems with this inference were identified from the experimental methodology used. In the initial exclusion test, Markman and Abelev (2004) suggested that both Rico and Chaser could have shown a base-line novelty preference for the new object, and retrieved it without learning the label used to refer to it. Obtaining a reward for choosing the new object may have influenced the subsequent retention test, as when presented with the new object plus four familiar and four unfamiliar objects, the dogs may have fetched the new object again as it had been previously rewarded, whilst the labels of the four familiar objects were already known and the four unfamiliar objects had never been rewarded. Griebel and Oller (2012) investigated this possibility using a Yorkshire terrier named Bailey. After successfully showing the ability to learn a range of word labels, Bailey was given the same exclusion and retention tests as used in the previous studies,

although in this case using two novel objects, and was similarly successful. She was then given an additional two-choice identification test, where both novel objects were presented together with no other objects. Bailey failed to retrieve the correct object in this task, demonstrating that she had not learnt the labels that referred to the new objects. This suggests that her success in the retention trial was through an extended form of exclusion learning, rather than showing evidence of 'fast mapping'. Therefore, it remains unclear if dogs can understand that words are symbolically referential, rather than learning that words match specific objects through associative learning. To determine more conclusively whether dogs perceive words as representational, researchers could employ a testing paradigm used to show this ability in bottlenosed dolphins (*Tursiops truncatus*) by establishing that they could understand reference to an absent object. Herman and Forestell (1985) trained a captive dolphin named Ake to successfully give 'yes' or 'no' responses by pressing different paddles when asked if specific objects were present in her tank. If dogs also understand that labels can refer to absent objects, this would demonstrate that they perceive words as abstract symbols for objects, rather than purely associating the word and object together.

As well as dogs' potentially using different cognitive mechanisms to humans, it is interesting to note that the associations dogs form to match word labels to the correct objects may be based on different object features to those one might expect. Humans tend to generalise a word that refers to one object to new objects based on their similarity in shape. In contrast, dogs may focus on different attributes, such as the size and texture of objects. This was demonstrated by another border collie (Gable), who had previous training in learning object labels. After initial training to fetch a novel object using an arbitrary word label (e.g., 'dax'), Gable was asked to fetch the 'dax' object and was presented a choice of two new objects, both of which differed from the original in size, shape, or texture. In similar experiments, humans tend to choose the new object with the same shape as the original (Landau et al., 1988), whereas Gable initially chose objects of the same size and with further familiarisation chose objects that had the same texture as the original object (Van der Zee et al., 2012). This could, of course, be linked to the fact that dogs frequently manipulate objects in their mouths, but also brings to light intriguing possibilities of potential differences to humans in the way that dogs may perceive object features

Finally, as well as forming concepts based on individual words or simple combinations, humans meaningfully arrange words using syntactic or grammatical rules, giving a limitless generative capacity to express and combine different concepts. By exploring how dogs respond to grammatical rules, researchers gain insights into the extent to which dogs share these cognitive abilities. To explore this, Pilley (2013) recently adapted a testing paradigm originally developed to demonstrate that bottlenosed dolphins could successfully respond to different relational combinations of known words. In his study, Pilley (2013) successfully taught Chaser to respond to sentences including three elements

of grammar: a prepositional object, a verb, and a direct object (e.g., 'to ball take frizbee'). In itself, this ability shows an impressive capacity in the working memory, but even more interestingly, Chaser responded correctly when the object labels were reversed, illustrating her awareness of word order and simple grammatical rules. As well as sharing this ability with bottlenosed dolphins, non-human primates have also been shown to process rule-governed sequences in acoustic signals [bonobo (*Pan paniscus*): Savage-Rumbaugh et al., 1993; cotton-top tamarins (*Saguinus oedipus*): Fitch & Hauser, 2004]. However, although dogs show comparable abilities to other language-trained species in spontaneously learning simple grammatical rules, known as finite state grammars, it is not clear whether they are able to process more complicated phrase structures. For example, although cotton-top tamarins also respond correctly to finite state grammars, they appear unable to spontaneously process phrase structures where related elements are distantly placed (e.g., 'if' followed by 'then') (Fitch & Hauser, 2004). It has yet to be established if dogs are able to learn these higher-order phrase-structure grammars.

Thus, despite being relatively constrained in their own vocal production, dogs have shown the ability to adapt their perceptual abilities in order to discriminate the phonetic elements of human speech. Whilst dogs show comparable comprehension skills to other language-trained animal species, we still have much to learn about the cognitive mechanisms underlying their speech perception, particularly how speech-related information is represented by dogs. Further to this, another interesting avenue to be explored is the extent to which dogs partition linguistic information from prosodic cues, and if this mirrors the way in which humans separate this information. Overall, it certainly seems that as well as responding exceptionally well to human spoken commands via label learning, dogs can also extract information about the speaker, allowing them to identify familiar individuals and discriminate some of their physical attributes. Dogs may furthermore recognise some level of emotion in the human voice, facilitated perhaps by dog owners' automatic use of exaggerated prosody. Much more research is needed to determine exactly how dogs learn to respond to cues in the human voice, particularly the extent to which interspecific communication can be explained by similarities across our species' vocal signals.

5.6 CONCLUSIONS

Throughout this chapter, we have seen that at a superficial level domestic dogs appear to possess a relatively limited vocal repertoire with fewer functional distinctions between call types than their wild ancestors. This initially appeared to indicate functional atrophy as a result of relaxed natural selection pressures during domestication. However, on closer examination of the acoustic structure of dog vocalisations, it is evident that their vocal repertoire has the ability to

broadcast a range of socially relevant information about the signaller, including his/her body size, motivational state, and almost certainly some measure of individual identity. Indeed, rather than indicating a functional decline, some differences in vocal production, as we have seen, appear to have become adapted to facilitate communication with humans.

Despite ongoing research, there remain many aspects of dog vocal communication that we do not yet fully understand, from the broadcasting of dynamic attributes such as motivational state within individual call types to the encoding of static attributes such as individual identity across different calls. It is also important to note that many studies to date have focussed on a single breed (e.g., Hungarian mudi; Pongrácz et al., 2005, 2006; Marós et al., 2008) or a very small sample resulting in a limited number of breeds (e.g., ten dogs; Yin, 2002; Yin & McCowan, 2004). Because the range of morphological and behavioural variation between different dog breeds is so great, this might have led to biased representations of the occurrence and communicative content of dog vocalisations. Using a wide range of different breeds in research can ensure that such potential differences are controlled for, thus promoting greater generalisability across domestic dogs as a species.

Although breed differences are undoubtedly present, at a species level, it is clear that acoustic communication plays a major role during social interactions, allowing domestic dogs to transmit a broad range of important social information. Recent experiments have greatly advanced not only our understanding of the physiological, perceptual, and cognitive processes underlying the dog's vocal communication system, but also of how some of these mechanisms may have adapted to facilitate interaction with humans during the process of domestication. Further research remains crucial to provide deeper insights into the full communicative potential of dog vocalisations during social interactions, both within and across species.

Future Directions

- Are motivational state and emotional valence encoded within call types?
- Do vocal signatures across call types in dogs support individual recognition?
- What is the function of the full range of acoustically diverse call types within and between breeds?
- Do dogs perceive and make use of emotional and motivational cues in human voices?
- How is acoustic information integrated with information from other modalities?
- Are dog vocal production and auditory perception adapted to facilitate communication with humans?

REFERENCES

Adachi, I., Fujita, K., 2007. Cross-modal representation of human caretakers in squirrel monkeys. Behav. Processes 74 (1), 27–32.

Adachi, I., Kuwahata, H., Fujita, K., 2007. Dogs recall their owner's face upon hearing their owner's voice. Anim. Cogn. 10 (1), 17–21.

Archer, J., Monton, S., 2011. Preferences for infant facial features in pet dogs and cats. Ethology 117 (3), 217–226.

Ashdown, R.R., Lea, T., 1979. The larynx of the basenji dog. J. Small Anim. Pract. 20 (11), 675–679.

Aubergé, V., Cathiard, M., 2003. Can we hear the prosody of smile? Speech Comm. 40, 87–97.

August, P.V., Anderson, J.G., 1987. Mammal sounds and motivation-structural rules: a test of the hypothesis. J. Mammal., 1–9.

Bachorowski, J.A., Owren, M.J., 1999. Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. J. Acoustical Soc. Am. 106, 1054.

Bachorowski, J., Owren, M.J., 2008. Vocal expressions of emotion. In: Lewis, M., Haviland-Jones, J.M., Feldman, L. (Eds.), Handbook of emotions, 3rd ed. Guilford Press, New York, pp. 196–210.

Bahrick, L.E., Hernandez-Reif, M., Flom, R., 2005. The development of infant learning about specific face-voice relations. Develop. Psychol. 41 (3), 541.

Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. J. Pers. Soc. Psychol. 70 (3), 614.

Baru, A.V., 1975. Discrimination of synthesized vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration and number of formants) in dog. Aud. Anal. Percept. Speech, 91–101.

Bekoff, M., 1974. Social play and play-soliciting by infant canids. Am. Zoologist 14 (1), 323–340.

Bekoff, M., Diamond, J., Mitton, J.B., 1981. Life-history patterns and sociality in canids: body size, reproduction, and behavior. Oecologia 50 (3), 386–390.

Bloom, P., 2004. Can a dog learn a word? Science 304 (5677), 1605–1606.

Bradshaw, J., Nott, H., 1995. Social and communication behaviour of the dog. In: Serpell, J. (Ed.), The domestic dog: its evolution, behaviour, and interactions with people. Cambridge University Press, Cambridge, pp. 116–130.

Brady, C.A., 1981. The vocal repertoires of the bush dog (*Speothos venaticus*), crab-eating fox (*Cerdocyon thous*), and maned wolf (*Chrysocyon brachyurus*). Anim. Behav. 29 (3), 649–669.

Bryant, G.A., Barrett, H.C., 2007. Recognizing intentions in infant-directed speech evidence for universals. Psychol. Sci. 18 (8), 746–751.

Burnham, D., Kitamura, C., Vollmer-Conna, U., 2002. What's new, pussycat? On talking to babies and animals. Science 296 (5572), 1435.

Buttelmann, D., Tomasello, M., 2012. Can domestic dogs (*Canis familiaris*) use referential emotional expressions to locate hidden food? Anim. Cogn. 16 (1), 137–145.

Cain, A.O., 1985. Pets as family members. Marriage Fam. Rev. 8 (3–4), 5–10.

Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., 2008. Encoding emotions in speech with the size code—a perceptual investigation. Phonetica 65, 210–230.

Cohen, J.A., Fox, M.W., 1976. Vocalizations in wild canids and possible effects of domestication. Behav. Processes 1 (1), 77–92.

Coleman, R.O., 1976. A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. Lang. Hear. Res. 14, 2–3.

Coppinger, R., Feinstein, M., 1991. Hark-hark, the dogs do bark and bark and bark—whether barking is an adaptive trait and has a purpose. Smithsonian 21 (10), 119.
Corbett, L.K., 1995. The dingo in Australia and Asia. Comstock/Cornell University Press, Ithaca, NY.
Coren, S., 2005. How to speak dog: the age of dog–human communication. Pocket Books, Simon and Schuster, London.
Custance, D., Mayer, J., 2012. Empathic-like responding by domestic dogs (*Canis familiaris*) to distress in humans: an exploratory study. Anim. Cogn. 15 (5), 851–859.
Deaux, E.C., Clarke, J.A., 2013. Dingo (*Canis lupus dingo*) acoustic repertoire: form and contexts. Behaviour 150 (1), 75–101.
Düpjan, S., Schön, P.C., Puppe, B., Tuchscherer, A., Manteuffel, G., 2008. Differential vocal responses to physical and mental stressors in domestic pigs (*Sus scrofa*). Appl. Anim. Behav. Sci. 114 (1), 105–115.
Estes, R.D., 1991. The behavior guide to African mammals: including hoofed mammals, carnivores, primates. University of California Press, Berkeley and Oxford.
Fant, G., 1970. Acoustic theory of speech production (No. 2). Walter de Gruyter, New York.
Faragó, T., Pongrácz, P., Miklósi, Á., Huber, L., Virányi, Z., Range, F., 2010. Dogs' expectation about signalers' body size by virtue of their growls. PLoS One 5 (12), e15175.
Fedderden-Petersen, D., 2004. Communication in wolves and dogs. In: Bekoff, M. (Ed.), Encyclopedia of animal behaviour, vol. 1. Garland SMTP Press, New York, pp. 385–394.
Feddersen-Petersen, D.U., 2000. Vocalization of European wolves (*Canis lupus lupus L.*) and various dog breeds (*Canis lupus f. fam.*). Archiv fur Tierzucht 43 (4), 387–398.
Fernald, A., 1989. Intonation and communicative intent in mothers' speech to infants: is the melody the message? Child Dev. 60 (6), 1497–1510.
Fernald, A., Kuhl, P., 1987. Acoustic determinants of infant directed preference for motherese speech. Infant Behav. Dev. 10 (3), 279–293.
Fitch, W.T., 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. J. Acoustical Soc. Am. 102, 1213.
Fitch, W.T., 2000a. The evolution of speech: a comparative review. Trends Cogn. Sci. 4 (7), 258–267.
Fitch, W.T., 2000b. The phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals. Phonetica 57 (2–4), 205–218.
Fitch, W.T., 2000c. Skull dimensions in relation to body size in nonhuman mammals: the causal bases for acoustic allometry. Zoology-Anal. Complex Syst. 103 (1–2), 40–58.
Fitch, W.T., Giedd, J., 1999. Morphology and development of the human vocal tract: a study using magnetic resonance imaging. J. Acoustical Soc. Am. 106 (3), 1511–1522.
Fitch, W.T., Hauser, M.D., 2004. Computational constraints on syntactic processing in nonhuman primate. Science 303 (5656), 377–380.
Fitch, W.T., Reby, D., 2001. The descended larynx is not uniquely human. Proc. Royal Soc. London. Series B: Biol. Sci. 268 (1477), 1669–1675.
Fox, M.W., 1984. Behaviour of wolves, dogs, and related canids. Krieger Publishing Co, Malabar, FL.
Frank, H., Frank, M.G., 1982. On the effects of domestication on canine social development and behavior. Appl. Anim. Ethol. 8 (6), 507–525.
Frynta, D., Baudyšová, J., Hradcová, P., Faltusová, K., Kratochvíl, L., 2012. Allometry of sexual size dimorphism in domestic dog. PloS One 7 (9), e46125.
Fukuzawa, M., Mills, D.S., Cooper, J.J., 2005. The effect of human command phonetic characteristics on auditory cognition in dogs (*Canis familiaris*). J. Comp. Psychol. 119 (1), 117.

Ghazanfar, A.A., Turesson, H.K., Maier, J.X., van Dinther, R., Patterson, R.D., Logothetis, N.K., 2007. Vocal-tract resonances as indexical cues in rhesus monkeys. Curr. Biol. 17 (5), 425–430.

Gisiner, R., Schusterman, R.J., 1992. Sequence, syntax, and semantics: responses of a language-trained sea lion (*Zalophus californianus*) to novel sign combinations. J. Comp. Psychol. 106 (1), 78–91.

Gittleman, J.L., 1989. Carnivore behavior, ecology and evolution. Cornell University Press, New York.

Griebel, U., Oller, D.K., 2012. Vocabulary learning in a Yorkshire terrier: slow mapping of spoken words. PloS One 7 (2), e30182.

Hare, B., Tomasello, M., 1999. Domestic dogs (*Canis familiaris*) use human and conspecific cues to locate hidden food. Comp. Psychol. 113 (2), 173–177.

Harrington, F., Mech, L., 1978. Wolf vocalization. In: Hall, R.L., Sharp, H.S. (Eds.), Wolf and man: evolution in parallel. Academic Press, New York, pp. 109–132.

Hauser, M.D., 2000. The sound and the fury: primate vocalizations as reflections of emotion and thought. In: Wallin, N.L., Merker, B., Brown, S. (Eds.), The origins of music. The MIT Press, Cambridge, MA, pp. 77–102.

Herman, L.M., Forestell, P.H., 1985. Reporting presence and absence of named objects by a language-trained dolphin. Neurosci. Behav. Rev. 9 (4), 667–681.

Herrel, A., Gibb, A.C., 2006. Ontogeny of performance in vertebrates. Physiol. Biochem. Zool. 79 (1), 1–6.

Hillis, T.L., Mallory, F.F., 1996. Sexual dimorphism in wolves (*Canis lupus*) of the Keewatin District, Northwest Territories, Canada. Can. J. Zool. 74 (4), 721–725.

Hirsh-Pasek, K., Treiman, R., 1982. Doggerel: motherese in a new context. J. Child Lang. 9 (1), 229–237.

Jaeger, M.M., Pandit, R.K., Haque, E., 1996. Seasonal differences in territorial behavior by golden jackals in Bangladesh: howling versus confrontation. J. Mammal, 768–775.

Joslin, P.W.B., 1967. Movements and home sites of timber wolves in Algonquin Park. Am. Zoologist 7 (2), 279–288.

Kaminski, J., Call, J., Fischer, J., 2004. Word learning in a domestic dog: evidence for 'fast mapping'. Science 304 (5677), 1682–1683.

Koler-Matznick, J., Brisbin, I.L., Feinstein, M., Bulmer, S., 2003. An updated description of the New Guinea singing dog (*Canis hallstromi*). J. Zool. 261 (2), 109–118.

Landau, B., Smith, L.B., Jones, S.S., 1988. The importance of shape in early lexical learning. Cogn. Dev. 3, 299–321.

Leaver, S.D.A., Reimchen, T.E., 2008. Behavioural responses of *Canis familiaris* to different tail lengths of a remotely-controlled life-size dog replica. Behaviour 145 (3), 377–390.

Lieberman, P., Blumstein, S., 1988. Speech physiology, speech perception, and acoustic phonetics. Cambridge University Press, Cambridge.

MacNulty, D.R., Smith, D.W., Mech, L.D., Eberly, L.E., 2009. Body size and predatory performance in wolves: is bigger better? J. Anim. Ecol. 78 (3), 532–539.

Markman, E.M., Abelev, M., 2004. Word learning in dogs? Trends Cogn. Sci. 8 (11), 479–481.

Marós, K., Pongrácz, P., Bárdos, G., Molnár, C., Faragó, T., Miklósi, A., 2008. Dogs can discriminate barks from different situations. Appl. Anim. Behav. Sci. 114, 159–167.

Mazzini, F., Townsend, S.W., Virányi, Z., Range, F., 2013. Wolf howling is mediated by relationship quality rather than underlying emotional stress. Curr. Biol. 23 (17), 1677–1680.

McComb, K., Taylor, A.M., Charlton, B., Wilson, C., 2009. The cry embedded within the purr. Curr. Biol. 19, 507–508.

McConnell, P.B., 1990. Acoustic structure and receiver response in domestic dogs, *Canis familiaris*. Anim. Behav. 39 (5), 897–904.

Merola, I., Prato-Previde, E., Lazzaroni, M., Marshall-Pescini, S., 2013. Dogs' comprehension of referential emotional expressions: familiar people and familiar emotions are easier. Anim. Cogn. 2013 Aug 18. [Epub ahead of print].

Miklósi, Á., Polgárdi, R., Topál, J., Csányi, V., 1998. Use of experimenter-given cues in domestic dogs. Anim. Cogn. 1, 113–121.

Miles, H.L., Harper, S.E., 1994. 'Ape language' studies and human language origins. Hominid Cult. in Primate Perspect., 253–278.

Mills, D.S., Fukuzawa, M., Cooper, J.J., 2005. The effect of emotional content of verbal commands on the response of dogs. In: Current issues and research in veterinary behavioural medicine. Purdue University Press, West Lafayette, IN (pp. 207–220). Papers presented at the 5th International Veterinary Behavior Meeting.

Moehlman, P.D., 1983. Socioecology of silverbacked and golden jackals (*Canis mesomelas* and *Canis aureus*). Adv. Study Mammal. Behav. 7, 423–453.

Molnár, C., Kaplan, F., Roy, P., Pachet, F., Pongrácz, P., Dóka, A., Miklósi, Á., 2008. Classification of dog barks: a machine learning approach. Anim. Cogn. 11, 389–400.

Molnár, C., Pongrácz, P., Dóka, A., Miklósi, A., 2006. Can humans discriminate between dogs on the base of the acoustic parameters of barks? Behav. Proces. 73, 76–83.

Molnár, C., Pongrácz, P., Faragó, T., Dóka, A., Miklósi, Á., 2009. Dogs discriminate between barks: the effect of context and identity of the caller. Behav. Processes 82, 198–201.

Morton, E.S., 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. Am. Nat., 855–869.

Ohala, J., 1984. An ethological perspective on common cross-language utilization of F0 of voice. Phonetica 41, 1–16.

Ohala, J., 1996. Ethological theory and the expression of emotion in the voice. In: Proc. ICSLP 96, [4th International Conference on Spoken Language Processing, Philadelphia], 3–6 October, 1996, Vol. 3, University of Delaware, Wilmington, DE, pp. 1812–1815. Accessed online 12 June, 2008. http://www.linguistics.berkeley.edu/phonlab/users/ohala/papers/emotion_in_the_voice.pdf.

Owings, D., Morton, E., 1998. Animal vocal communication: a new approach. Cambridge University Press, Cambridge.

Owren, M.J., Rendall, D., 1997. An affect-conditioning model of nonhuman primate vocal signalling. Perspect. Ethol. 12, 299–346.

Owren, M.J., Rendall, D., 2001. Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signalling. Evolutionary Anthropol.: Issues, News, Rev. 10 (2), 58–71.

Pepperberg, I.M., 1981. Functional vocalizations by an African grey parrot (*Psittacus erithacus*). Ethology 55 (2), 139–160.

Peters, R.H., 1986. The ecological implications of body size. Cambridge University Press, Cambridge.

Piérard, J.A.M., 1963. Comparative anatomy of the carnivore larynx: with special reference to the cartilages and muscles of the larynx in the dog. Cornell University, New York.

Pilley, J.W., 2013. Border collie comprehends sentences containing prepositional object, verb, and direct object. Learn. Motiv. 44 (4), 229–240.

Pilley, J.W., Reid, A.K., 2011. Border collie comprehends object names as verbal referents. Behav. Processes 86 (2), 184–195.

Plotsky, K., Rendall, D., Riede, T., Chase, K., 2013. Radiographic analysis of vocal tract length and its relation to overall body size in two canid species. J. Zool. 291 (1), 76–86.

Pongrácz, P., Molnár, C., Miklósi, Á., 2006. Acoustic parameters of dog barks carry emotional information for humans. Appl. Anim. Behav. Sci. 100 (3), 228–240.

Pongrácz, P., Molnár, C., Miklósi, Á., Csányi, V., 2005. Human listeners are able to classify dog (*Canis familiaris*) barks recorded in different situations. J. Comp. Psychol. 119 (2), 136.

Prato-Previde, E., Fallani, G., Valsecchi, P., 2006. Gender differences in owners interacting with pet dogs: an observational study. Ethology 112 (1), 64–73.

Price, E.O., 1999. Behavioral development in animals undergoing domestication. Appl. Anim. Behav. Sci. 65 (3), 245–271.

Proops, L., McComb, K., Reby, D., 2009. Cross-modal individual recognition in domestic horses (*Equus caballus*). Proc. Natl. Acad. Sci. 106 (3), 947–951.

Puts, D.A., Gaulin, S.J., Verdolini, K., 2006. Dominance and the evolution of sexual dimorphism in human voice pitch. Evolution Hum. Behav. 27 (4), 283–296.

Puts, D.A., Hodges, C.R., Cárdenas, R.A., Gaulin, S.J., 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. Evolution Hum. Behav. 28 (5), 340–344.

Ramos, D., Ades, C., 2012. Two-item sentence comprehension by a dog (*Canis familiaris*). Plos One 7 (2), e29689.

Ratcliffe, V.F., McComb, K., & Reby, D. (submitted). Effect of experience on domestic dogs' (*Canis familiaris*) ability to spontaneously associate the gender of unfamiliar humans across modalities.

Reby, D., McComb, K., 2003. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. Anim. Behav. 65, 519–530.

Rendall, D., 2003. The affective basis of referential grunt variants in baboons. J. Acoustical Soc. Am. 113, 3390–3402.

Rendall, D., Kollias, S., Ney, C., Lloyd, P., 2005. Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. J. Acoustical Soc. Am. 117, 944–955.

Riede, T., Fitch, T., 1999. Vocal tract length and acoustics of vocalisation in the domestic dog (*Canis familiaris*). J. Exp. Biol. 202, 2859–2867.

Robbins, R.L., 2000. Vocal communication in free-ranging African wild dogs (*Lycaon pictus*). Behaviour 137 (10), 1271–1298.

Rutter, R., Pimlott, D., 1968. The world of the wolf. Lippincott, Philadelphia, PA.

Ryalls, J.H., Lieberman, P., 1982. Fundamental-frequency and vowel perception. J. Acoustical Soc. Am. 72, 1631–1634.

Sauter, D.A., Eisner, F., Ekman, P., Scott, S.K., 2009. Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. Proc. Natl Acad. Sci. 107 (6), 2408–2412.

Savage-Rumbaugh, E.S., Murphy, J., Sevcik, R.A., Brakke, K.E., Williams, S.L., Rumbaugh, D.M., Bates, E., 1993. Language comprehension in ape and child. Monogr. Soc. Res. Child Dev. i-252.

Schassburger, R.M., 1993. Vocal communication in the timber wolf, Canis lupus: structure, motivation, and ontogeny. Paul Parey Scientific Publishers, Berlin.

Scheider, L., Grassmann, S., Kaminski, J., Tomasello, M., 2011. Domestic dogs use contextual information and tone of voice when following a human pointing gesture. Plos One 6 (7), e21676.

Schmidt-Nielsen, K., 1975. Scaling in biology: the consequences of size. J. Exp. Zool. 194 (1), 287–307.

Shamir, M.H., Leisner, S., Klement, E., Gonen, E., Johnston, D.E., 2002. Dog bite wounds in dogs and cats: a retrospective study of 196 cases. J. Vet. Med. Series A. 49 (2), 107–112.
Sillero-Zubiri, C., Gottelli, D., 1994. Canis simensis. Mammalian Species 485, 1–6.
Sillero-Zubiri, C., Hoffmann, M., Macdonald, D.D.W. (Eds.), 2004. Canids: foxes, wolves, jackals and dogs: status survey and conservation action plan, vol. 62. IUCN, Oxford.
Smith, D., Patterson, R., 2005. The interaction of glottal pulse rate and vocal tract length in judgement of speaker's size, sex and age. J. Acoustical Soc. Am. 118, 3177–3186.
Taylor, A.M., Reby, D., 2010. The contribution of the source–filter theory to mammal vocal communication. J. Zool. 280 (3), 221–236.
Taylor, A.M., Reby, D., McComb, K., 2008. Human listeners attend to size information in domestic dog growls. J. Acoustical Soc. Am. 123, 2903.
Taylor, A.M., Reby, D., McComb, K., 2009. Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). Ethology 115 (10), 905–915.
Taylor, A.M., Reby, D., McComb, K., 2010a. Size communication in domestic dog *(Canis familiaris)* growls. Anim. Behav. 79, 205–210.
Taylor, A.M., Reby, D., McComb, K., 2010b. Why do large dogs sound more aggressive to human listeners: acoustic bases of motivational misattributions. Ethology 116, 1155–1162.
Taylor, A.M., Reby, D., McComb, K., 2011. Cross modal perception of body size in domestic dogs *(Canis familiaris)*. PLoS One 6 (2), e17069.
Theberge, J.A., Falls, J.B., 1967. Howling as a means of communication in timber wolves. Am. Zoologist 7 (2), 331–338.
Titze, I.R., 1994. Principles of vocal production. Prentice Hall, Englewood Cliffs, NJ.
Titze, I.R., 2000. Principles of voice production. National Centre for Voice and Speech, Iowa City, IA.
Van der Zee, E., Zulch, H., Mills, D., 2012. Word generalization by a dog (*Canis familiaris*): is shape important? Plos One 7 (11), e49382.
Vitulli, W.F., 2006. Attitudes towards empathy in domestic dogs and cats 1, 2. Psychol. Rep. 99 (3), 981–991.
Volodin, I.A., Volodina, E.V., Isaeva, I.V., 2001. Vocal repertoire in the dhole *Cuon alpinus* (Carnivora, Canidae) in captivity. Entomol. Rev. 81 (2), S346.
Wayne, R.K., 1993. Molecular evolution of the dog family. Trends Genetics 9 (6), 218–224.
Yin, S., 2002. A new perspective on barking in dogs (*Canis familiaris*). J. Comp. Psychol 116, 189–193.
Yin, S., McCowan, B., 2004. Barking in domestic dogs: context specificity and individual identification. Anim. Behav. 68, 343–355.
Zaccaroni, M., Passilongo, D., Buccianti, A., Dessì-Fulgheri, F., Facchini, C., Gazzola, A., Apollonio, M., 2012. Group specific vocal signature in free-ranging wolf packs. Ethol. Ecol. Evol. 24 (4), 322–331.
Zuberbuhler, K., 2002. A syntactic rule in forest monkey communication. Anim. Behav. 63, 293–299.